(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 June 2002 (13.06.2002)

PCT

(10) International Publication Number
WO 02/45523 A2

(51) International Patent Classification[7]: A23J 3/34

(21) International Application Number: PCT/EP01/14479

(22) International Filing Date: 6 December 2001 (06.12.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
00204404.8 7 December 2000 (07.12.2000) EP
01204464.0 15 November 2001 (15.11.2001) EP

(71) Applicant *(for all designated States except US)*: **DSM N.V.** [NL/NL]; Het Overloon 1, NL-6411 Te Heerlen (NL).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **EDENS, Luppo** [NL/NL]; Hoflaan 118, NL-3062 JL Rotterdam (NL). **HOEVEN, VAN DER, Robertus, Antonius, Mijndert** [NL/NL]; Kerkstraat 6a, NL-2223 AJ Katwijk (NL). **DELEST, Veronique** [FR/FR]; 53 Rue Du Chateau, F-92100 Boulogne (FR).

(74) Agents: **MATULEWICZ, Emil, Rudolf, Antoni** et al.; DSM N.V., DSM Patents & Trademarks, P.O. Box 1, NL-2600 MA Delft (NL).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*
— *entirely in electronic form (except for this front page) and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/45523 A2

(54) Title: PROTEIN HYDROLYSATES ENRICHED IN PEPTIDES HAVING A CARBOXY TERMINAL PROLINE RESIDUE

(57) Abstract: A method of enzymatically producing a protein hydrolysate from a protein substrate is described, wherein a proline-specific endoprotease or a composition containing a proline-specific endoprotease and optionally a subtilisin or a metallo endoprotease, and other enzymes such as carboxypeptidases, is used to produce a protein hydrolysate enriched in peptide fragments having a carboxy terminal proline residue. Such protein hydrolysates may be used as such or to reduce bitterness in foods nutritionally supplemented by protein hydrolysates, as well as to produce hydrolysate-containing foodstuffs having low antigenicity.

indispensable step of removing these partially digested protein fragments from the hydrolysate inevitably lowers the yield of the final digestion product, thereby increasing production costs.

Protein antigenicity may be overcome by digesting proteins to peptides having only 8-10 amino acid residues, but the peptides created by such an extensive proteolytic digestion can be very bitter. The general explanation for this phenomenon is that smaller peptides with a high content of hydrophobic amino acids promote bitter tastes. The nature of the proteinaceous raw material used, the type of proteolytic enzymes used for digestion and the length of peptides obtained largely determine the degree of bitterness associated with the final hydrolysate. For example, casein, which contains many hydrophobic amino acids, is known to generate far more bitter hydrolysates than whey proteins.

In industrial operations, debittering of protein hydrolysates is carried out by the selective removal of bitter peptides using activated carbon or adsorption to hydrophobic resin. The concomitant yield reduction during such removal steps increases the cost of the final product. Moreover, this process has a negative impact on the nutritional value of the final product, as several nutritionally indispensable amino acids may be lost due to their hydrophobic nature, including tryptophan, leucine, phenylalanine and isoleucine. Thus, debittering in this way is prone to producing hydrolysates deficient in these nutritionally important amino acids.

Debittering can also be achieved by subjecting hydrolysates to exopeptidases. In this approach, amino-terminal and carboxy-terminal amino acids are liberated from peptides in an attempt to reduce their overall hydrophobicity. Exposure of peptides to non-selective exoproteases unfortunately results in the release of uncontrollable quantities of free amino acids into the final hydrolysate. Subsequent heating of such hydrolysates containing free amino acids, as required for sterilisation or spray drying, often generates brothy off-flavours via Maillard reactions. Moreover, the high levels of free amino acids created by exoproteases may increase the osmotic value of the final hydrolysate product to levels that can cause osmotic diarrhoea.

Therefore, the production of protein hydrolysates represents a trade-off between the pros and cons of proteolytic digestion. Current practise is to optimize enzymatic

Because of the negative influence on product cost and quality associated with protein hydrolysis, several enzyme mixtures aimed at improving hydrolysate characteristics and lowering production costs have been described in prior publications. Examples include EP 321 603, which refers to the use of animal-derived endoproteases like trypsin, chymotrypsin and pancreatin, and EP 325 986 and WO 96/13174, which favor the use of endoproteases obtained from *Bacillus* or *Aspergillus* species. Several exoproteases have been described as being capable of debittering mixtures of peptides. Whereas, for example, EP 0223 560 refers to the use of a specific proline specific endoprotease, WO 96/13174 refers to a mixture of amino-peptidases and carboxypeptidases for this purpose.

A number of publications tout the beneficial effects of proline-specific endoproteases in combination with various exopeptidases for producing protein hydrolysates which have relatively low bitterness profiles. For example, Japanese patent JP02039896 refers to the use of a proline-specific endoprotease combined with a dipeptidyl-carboxypeptidase for generating low molecular weight peptide preparations. The degradation of proline-rich oligopeptides by three proline-specific peptide hydrolases is described as essential for accelerating cheese ripening without bitterness (Journal of Dairy Science, 77 (2) 385-392 (1994)). More specifically, the debittering effect of proline-specific endoprotease in combination with a carboxypeptidase is described in JP5015314. JP5015314 describes a crude enzyme preparation obtained from *Aspergillus oryzae* that exhibits, apart from a general, non-specific proteolytic activity, small quantities of a proline-specific endoprotease and carboxypeptidase activity. According to JP5015314, proline residues present at the carboxy terminii of peptides cause bitter tastes and are undesirable. Incubation of soy bean protein with a proline-specific endoprotease and carboxypeptidase enzyme mixture yielded a hydrolysate that was significantly less bitter than a soy bean hydolysate obtained with protease preparation lacking the combination of a proline-specific endoprotease and a carboxypeptidase.

Collectively, the state of the art strongly suggests that exopeptidase-mediated release of carboxy terminal (or amino terminal) hydrophobic amino acid residues from peptides is essential for significantly debittering peptide hydrolysates. Likewise, references that specifically refer to proline-specific endoproteases for debittering teach that the function

preferably over 100 nucleotides, more preferably at least 90% identical over 200 nucleotides, or (ii) a nucleic acid sequence complementary to the nucleic acid sequence of SEQ ID NO: 1.

and a DNA molecule encoding the endopeptidase.

The present invention also provides:

- the use of a protein hydrolysate of the invention in a food or drink;
- the use of a proline-specific endoprotease according to the invention;
- a method of enzymatically producing a protein hydrolysate from a protein substrate, wherein the protein substrate is incubated with a proline-specific endoprotease to produce a protein hydrolysate enriched in peptides having a carboxy terminal proline;
- an enzyme composition comprising a proline-specific endoprotease of the invention, the composition being capable of producing a protein hydrolysate comprising peptides, wherein the molar fraction of peptides (%) carrying a carboxy terminal proline is at least two times the molar fraction (%) of proline in the protein or a hydrolysate of the invention; and
- a food comprising a protein hydrolysate of the invention or obtainable by a method of the invention.

Detailed description of the Invention

We have shown that a high incidence of proline residues at the carboxy terminal end of peptides can be correlated with low bitterness. Moreover we have demonstrated that the desired high incidence of carboxy terminal proline residues can only be achieved with high concentrations of a proline-specific endoprotease , i.e. concentrations that exceed the activity specified in JP5015314 by several orders of magnitude and moreover in the absence of a carboxypeptidase.

From an economic point of view the implication of this observation is that there exists a clear need for an improved means of producing proline-specific endoproteases in high quantities and a relatively pure form. A preferred way of doing this is via the overproduction of such a proline-specific endoprotease using recombinant DNA techniques. As many food products are acidic and long term enzyme incubations under industrial, non-sterile circumstances require acidic incubation conditions to prevent

the use of the proline-specific endoprotease to generate proline-rich peptides. Such proline-rich peptides are desirable additions to various food or nutraceutical products as they have been implicated in anorectic action, in fibrinolytic and antithrombotic and antihypertensive effects, in protection of the gastric mucosa as well as the prevention of rheumatoid arthritis.

Another surprising application is addition of the new enzyme to animal feed to enhance protein utilisation. For example, addition of the enzyme leads to improved digestibility of hard-to-digest proline rich sequences present in the feed protein as well as to improved conversion rates of cheaply available vegetable proteins containing high levels of polyphenols.

In yet another new application the enzyme is used in beer brewing. Barley proteins are rich in proline rich sequences and in their non-malted form cereal proteins are extremely difficult to degrade into the free amino acids required to create a suitable fermentable wort. Quite surprisingly the incorporation of the new enzyme into the mashing process has been shown to stimulate amino acid release from milled but non-malted barley so that a much richer wort is obtained. In a similar way beer fermentation from mashes containing a high proportion of other cheap and locally available cereals such as for example sorghum can be improved.

In most of these new applications the proline-specific endoprotease should preferably exhibit an activity spectrum with an acidic pH optimum.

To overcome the above-mentioned problems, the invention demonstrates that the activity of an isolated, purified proline-specific endoprotease alone, i.e. without the substantial concomitant or subsequent activity of an exoproteolytic enzyme, is sufficient for significantly debittering a protein hydrolysate. Therefore the proline-specific endoprotease may comprise at least 5 units per gram protein of the enzyme preparation of the invention, preferably 10 u/g, more preferably 25 u/g and even more preferably 50 u/g Moreover, studies conducted in accordance with the invention demonstrate that the activity of an isolated, purified proline-specific endoprotease alone, meaning without the concomitant or subsequent activity of an exoproteolytic enzyme, is sufficient to significantly decrease the overall immunogenicity level of protein hydrolysates, as well as to significantly increase their overall solubility under acidic conditions. The hydrolysates produced according to the invention are enriched in peptides having a carboxy terminal proline residue.

endoprotease or a metallo endoprotease and a carboxypeptidase to produce a protein hydrolysate enriched in peptide fragments having a carboxy terminal proline.

The enzyme mixture of the invention is particularly suitable for use in the production of protein hydrolysates intended for flavoring and nutrient enhancement of sport drinks and juice-based beverages, as the resulting hydrolyzed peptide mixture combines a very low bitterness profile with excellent solubility under the prevailing acidic conditions of such beverages. The enzyme mixture of the invention is characterised in that it contains at least one endoprotease for example a serine protease or a metallo endoprotease in conjunction with a proline-specific endoprotease (E.C. 3.4..21.26) to provide a primary hydrolysate. More specifically, the invention relates to an isolated, purified proline-specific endoprotease and a serine protease or metallo protease enzyme mixture capable of producing a protein hydrolysate comprising peptide fragments, wherein at least 8%, preferably at least 15%, more preferably from 30 to 70% of said peptide fragments have a carboxy terminal proline.

Another embodiment of the invention is a protein hydrolysate enriched with a relatively high content of peptides having proline as the carboxy terminal amino acid residue. Such enriched hydrolysates may comprise at least 8%, preferably at least 15%, more preferably from 30 to 70% peptide fragments having a carboxy terminal proline residue. Since enzyme preparations typically utilized in the genesis of protein hydrolysates are not capable of generating peptides bearing proline residues at carboxy terminii, protein hydrolysates that are relatively rich in such peptides are novel.

Substrates for hydrolysis by an enzyme mixture of the invention include whole milk, skimmed milk, acid casein, rennet casein, acid whey products or cheese whey products. Quite surprisingly the *Aspergillus* derived proline specific endoprotease doesnot only cleave at the carboxy-terminal side of proline residues but also at the carboxy-terminal side of hydroxyproline residues which makes other, collagen based animal proteins such as gelatine as well as bones or fish-bones containing residual meat, interesting substrates for the enzyme. Moreover, vegetable substrates like wheat gluten, milled barley and protein fractions obtained from, for example, soy, rice or corn are suitable substrates. Milk protein hydrolysates produced according to the invention may be used with or without additional filtration or purification steps in various speciality foods such as hypoallergenic hydrolysates for infant nutrition, basic hydrolysates for enteral and dietetic nutrition, as well as protein concentrates for various forms of health food. Thus, protein hydrolysates of the invention may be used to produce foodstuffs

hydrophobic amino acids such as Tyr, Trp, Phe and Leu. The enzyme mixture of the invention may contain chymotrypsin and/or subtilisin. Subtilisin is produced by species of Bacillus, has a particularly broad substrate specificity and a broad, alkaline pH optimum. The enzyme is optimally active between 50°C and 60°C. The enzyme is cheaply available as a regular commercial product and is useful in the production of, for example, various milk hydrolysates. Chymotrypsin may be obtained from animal pancreases, has a somewhat narrower substrate specificity at slightly more alkaline pH values than subtilisin and is optimally active below 50 degrees C.

The class of metallo endoproteases is wide spread in bacteria, fungi and higher organisms.They can be separated into the neutral and acid metalloproteases. Of these two subclasses only the neutral proteases exhibit the desirable cleavage preference i.e. cleaving the peptide chain on the carboxy terminal side of hydrophobic amino acid residues such as Phe and Leu. Well known representants of the category of the neutral metallo proteases are bacillolysin (E.C. 3.4.24.28) and thermolysin (E.C. 3.4.24.27) and either, or both of these, may be present in the enzyme mixture of the invention. Both enzymes are obtained from *Bacillus* species and exhibit maximum activity under neutral or slightly alkaline conditions. Less well known representants of these neutral metallo endoproteases have been obtained from *Aspergillus* species. In those cases in which the proline specific endoprotease is not used for its debittering effects but to aid in the hydrolysis of proline rich protein sequences, combinations with the class of the acid metalloproteases, as for example deuterolysine (EC 3.4.24.39) can be advantageous.

A proline-specific endoprotease is an endoprotease capable of cleaving peptides or polypeptides at the carboxy-terminal end of proline residues. Such enzymes are widely found in animals and plants, but their presence in microorganisms appears to be limited. To date, proline-specific endoprotease have been identified in species of *Aspergillus* (EP 0 522 428), *Flavobacterium* (EP 0 967 285) and *Aeromonas* (J.Biochem.113, 790-796), *Xanthomonas* and *Bacteroides*. Though the proline-specific enzymes from most of these organisms are active around pH 8, the *Aspergillus* enzyme is optimally active around pH 5. According to a preferred embodiment, proline-specific endoprotease having a pH optimum below 7, preferably having a pH optimum from 3.5 to 6.5 is used because of the technical and economical advantages of such enzymes. The proline-specific endoprotease of the invention may be isolated from one of the above-mentioned microbial species, particularly from a species of *Aspergillus*. Preferably, the proline-specific endoprotease is isolated from a strain of *Aspergillus niger*. More preferably, the

example more than 30%, 40%, 50%, 80%, 90%, 95% or 99%, by weight of the proteins in the preparation is a polypeptide of the invention.

Preferably, the polypeptide of the invention is obtainable from a microorganism which possesses a gene encoding an enzyme with proline specific endoprotease activity. More preferably the microorganism is fungal, and optimally is a filamentous fungus. Preferred organisms are thus of the genus *Aspergillus*, such as those of the species *Aspergillus niger*.

In a first embodiment, the present invention provides an isolated polypeptide having an amino acid sequence which has a degree of amino acid sequence identity to amino acids 1 to 526 of SEQ ID NO: 2 (i.e. the polypeptide) of at least about 40%, preferably at least about 50%, preferably at least about 60%, preferably at least about 65%, preferably at least about 70%, more preferably at least about 80%, even more preferably at least about 90%, still more preferably at least about 95%, and most preferably at least about 97%, and which has proline specific endoprotease activity.

For the purposes of the present invention, the degree of identity between two or more amino acid sequences is determined by BLAST P protein database search program (Altschul et al., 1997, Nucleic Acids Research 25: 3389-3402) with matrix Blosum 62 and an expected threshold of 10.

A polypeptide of the invention may comprise the amino acid sequence set forth in SEQ ID NO: 2 or a substantially homologous sequence, or a fragment of either sequence having proline specific endoprotease activity. In general, the naturally occurring amino acid sequence shown in SEQ ID NO: 2 is preferred.

The polypeptide of the invention may also comprise a naturally occurring variant or species homologue of the polypeptide of SEQ ID NO: 2.

A variant is a polypeptide that occurs naturally in, for example, fungal, bacterial, yeast or plant cells, the variant having proline specific endoprotease activity and a sequence substantially similar to the protein of SEQ ID NO: 2. The term "variants" refers to polypeptides which have the same essential character or basic biological functionality as the proline specific endoprotease of SEQ ID NO: 2, and includes allelic variants. The essential character of proline specific endoprotease of SEQ ID NO: 2 is that it is an enzyme capable of cleaving the amino-terminal amino acid from a protein or (poly)peptide. Preferably, a variant polypeptide has at least the same level of proline specific endoprotease activity as the polypeptide of SEQ ID NO: 2. Variants include

amino acids within the protein sequence, as long as the peptide maintains the basic biological functionality of the proline specific endoprotease of SEQ ID NO: 2.

Amino acid substitutions may be made, for example from 1, 2 or from 3 to 10, 20 or 30 substitutions. The modified polypeptide will generally retain activity as an proline specific endoprotease. Conservative substitutions may be made; such substitutions are well known in the art. Preferably substitutions do not affect the folding or activity of the polypeptide.

Shorter polypeptide sequences are within the scope of the invention. For example, a peptide of at least 50 amino acids or up to 60, 70, 80, 100, 150 or 200 amino acids in length is considered to fall within the scope of the invention as long as it demonstrates the basic biological functionality of the proline specific endoprotease of SEQ ID NO: 2. In particular, but not exclusively, this aspect of the invention encompasses the situation in which the protein is a fragment of the complete protein sequence.

In a second embodiment, the present invention provides an to isolated polypeptide which has proline specific endoprotease activity, and is encoded by polynucleotides which hybridize or are capable of hybrizing under low stringency conditions, more preferably medium stringency conditions, and most preferably high stringency conditions, with (I) the nucleic acid sequence of SEQ ID NO: 1 or a nucleic acid fragment comprising at least the c-terminal portion of SEQ ID NO: 1, but having less than all or having bases differing from the bases of SEQ ID NO: 1; or (ii) with a nucleic acid strand complementary to SEQ ID NO: 1.

The term "capable of hybridizing" means that the target polynucleotide of the invention can hybridize to the nucleic acid used as a probe (for example, the nucleotide sequence set forth in SEQ. ID NO: 1, or a fragment thereof, or the complement of SEQ ID NO: 1) at a level significantly above background. The invention also includes the polynucleotides that encode the proline specific endoprotease of the invention, as well as nucleotide sequences which are complementary thereto. The nucleotide sequence may be RNA or DNA, including genomic DNA, synthetic DNA or cDNA. Preferably, the nucleotide sequence is DNA and most preferably, a genomic DNA sequence. Typically, a polynucleotide of the invention comprises a contiguous sequence of nucleotides which is capable of hybridizing under selective conditions to the coding sequence or the complement of the coding sequence of SEQ ID NO: 1. Such nucleotides can be synthesized according to methods well known in the art.

be made which would result in a conservative amino acid substitution when the modified sequence is translated, for example as discussed with reference to polypeptides later.

Homologues

A nucleotide sequence which is capable of selectively hybridizing to the complement of the DNA coding sequence of SEQ ID NO: 1 is included in the invention and will generally have at least 50% or 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98% or at least 99% sequence identity to the coding sequence of SEQ ID NO: 1 over a region of at least 60, preferably at least 100, more preferably at least 200 contiguous nucleotides or most preferably over the full length of SEQ ID NO: 1. Likewise, a nucleotide which encodes an active proline specific endoprotease and which is capable of selectively hybridizing to a fragment of a complement of the DNA coding sequence of SEQ ID NO: 1, is also embraced by the invention. A C-terminal fragment of the nucleic acid sequence of SEQ ID NO:1 which is at least 80% or 90% identical over 60, preferably over 100 nucleotides, more preferably at least 90% identical over 200 nucleotides is encompassed by the invention.

Any combination of the above mentioned degrees of identity and minimum sizes may be used to define polynucleotides of the invention, with the more stringent combinations (i.e. higher identity over longer lengths) being preferred. Thus, for example, a polynucleotide which is at least 80% or 90% identical over 60, preferably over 100 nucleotides, forms one aspect of the invention, as does a polynucleotide which is at least 90% identical over 200 nucleotides.

The UWGCG Package provides the BESTFIT program which may be used to calculate identity (for example used on its default settings).

The PILEUP and BLAST N algorithms can also be used to calculate sequence identity or to line up sequences (such as identifying equivalent or corresponding sequences, for example on their default settings).

Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pair (HSPs) by identifying short words of length W in the query sequence that either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold. These initial neighborhood word hits act as seeds for initiating searches to find HSPs containing

from a yeast, bacterial, plant, prokaryotic or fungal cell, preferably of an *Aspergillus* strain, performing a polymerase chain reaction under conditions suitable for the amplification of the desired region, isolating the amplified fragment (e.g. by purifying the reaction mixture on an agarose gel) and recovering the amplified DNA. The primers may be designed to contain suitable restriction enzyme recognition sites so that the amplified DNA can be cloned into a suitable cloning vector.

Such techniques may be used to obtain all or part of the polynucleotides encoding the proline specific endoprotease sequences described herein. Introns, promoter and trailer regions are within the scope of the invention and may also be obtained in an analogous manner (e.g. by recombinant means, PCR or cloning techniques), starting with genomic DNA from a fungal, yeast, bacterial plant or prokaryotic cell.

The polynucleotides or primers may carry a revealing label. Suitable labels include radioisotopes such as ^{32}P or ^{35}S, enzyme labels, or other protein labels such as biotin. Such labels may be added to polynucleotides or primers of the invention and may be detected using techniques known to persons skilled in the art.

Polynucleotides or primers (or fragments thereof) labelled or unlabelled may be used in nucleic acid-based tests for detecting or sequencing an proline specific endoprotease or a variant thereof in a fungal sample. Such detection tests will generally comprise bringing a fungal sample suspected of containing the DNA of interest into contact with a probe comprising a polynucleotide or primer of the invention under hybridizing conditions, and detecting any duplex formed between the probe and nucleic acid in the sample. Detection may be achieved using techniques such as PCR or by immobilizing the probe on a solid support, removing any nucleic acid in the sample which is not hybridized to the probe, and then detecting any nucleic acid which is hybridized to the probe. Alternatively, the sample nucleic acid may be immobilized on a solid support, the probe hybridized and the amount of probe bound to such a support after the removal of any unbound probe detected.

The probes of the invention may conveniently be packaged in the form of a test kit in a suitable container. In such kits the probe may be bound to a solid support where the assay format for which the kit is designed requires such binding. The kit may also contain suitable reagents for treating the sample to be probed, hybridizing the probe to nucleic acid in the sample, control reagents, instructions, and the like. The probes and polynucleotides of the invention may also be used in microassay.

enzyme recognition sites, or to alter the property or function of the polypeptides encoded by the polynucleotides.

The invention includes double stranded polynucleotides comprising a polynucleotide of the invention and its complement.

The present invention also provides polynucleotides encoding the polypeptides of the invention described above. Since such polynucleotides will be useful as sequences for recombinant production of polypeptides of the invention, it is not necessary for them to be capable of hybridising to the sequence of SEQ ID NO: 1, although this will generally be desirable. Otherwise, such polynucleotides may be labelled, used, and made as described above if desired.

Recombinant Polynucleotides.

The invention also provides vectors comprising a polynucleotide of the invention, including cloning and expression vectors, and in another aspect methods of growing, transforming or transfecting such vectors into a suitable host cell, for example under conditions in which expression of a polypeptide of, or encoded by a sequence of, the invention occurs. Provided also are host cells comprising a polynucleotide or vector of the invention wherein the polynucleotide is heterologous to the genome of the host cell. The term "heterologous", usually with respect to the host cell, means that the polynucleotide does not naturally occur in the genome of the host cell or that the polypeptide is not naturally produced by that cell. Preferably, the host cell is a yeast cell, for example a yeast cell of the genus *Kluyveromyces* or *Saccharomyces* or a filamentous fungal cell, for example of the genus *Aspergillus*.

Polynucleotides of the invention can be incorporated into a recombinant replicable vector, for example a cloning or expression vector. The vector may be used to replicate the nucleic acid in a compatible host cell. Thus, in a further embodiment, the invention provides a method of making polynucleotides of the invention by introducing a polynucleotide of the invention into a replicable vector, introducing the vector into a compatible host cell, and growing the host cell under conditions which bring about replication of the vector. The vector may be recovered from the host cell. Suitable host cells are described below in connection with expression vectors.

Vectors

Preferred selectable markers include but are not limited to those that complement a defect in the host cell or confer resistance to a drug. They include for example versatile marker genes that can be used for transformation of most filamentous fungi and yeasts such as acetamidase genes or cDNAs (the amdS, niaD, facA genes or cDNAs from *A.nidulans*, *A.oryzae*, or *A.niger*), or genes providing resistance to antibiotics like G418, hygromycin, bleomycin, kanamycin, phleomycin or benomyl resistance (benA). Alternatively, specific selection markers can be used such as auxotrophic markers which require corresponding mutant host strains: e.g. URA3 (from *S.cerevisiae* or analogous genes from other yeasts), pyrG or pyrA (from *A.nidulans* or *A.niger*), argB (from *A.nidulans* or *A.niger*) or trpC. In a preferred embodiment the selection marker is deleted from the transformed host cell after introduction of the expression construct so as to obtain transformed host cells capable of producing the polypeptide which are free of selection marker genes.

Other markers include ATP synthetase subunit 9 (oliC), orotidine-5'-phosphate-decarboxylase (pvrA), the bacterial G418 resistance gene (useful in yeast, but not in filamentous fungi), the ampicillin resistance gene (*E. coli*), the neomycin resistance gene (*Bacillus*) and the *E. coli* uidA gene, coding for glucuronidase (GUS). Vectors may be used *in vitro*, for example for the production of RNA or to transfect or transform a host cell.

For most filamentous fungi and yeast, the expression construct is preferably integrated into the genome of the host cell in order to obtain stable transformants. However, for certain yeasts suitable episomal vector systems are also available into which the expression construct can be incorporated for stable and high level expression. Examples thereof include vectors derived from the 2 µm, CEN and pKD1 plasmids of *Saccharomyces* and *Kluyveromyces*, respectively, or vectors containing an AMA sequence (e.g. AMA1 from *Aspergillus*). When expression constructs are integrated into host cell genomes, the constructs are either integrated at random loci in the genome, or at predetermined target loci using homologous recombination, in which case the target loci preferably comprise a highly expressed gene. A highly expressed gene is a gene whose mRNA can make up at least 0.01% (w/w) of the total cellular mRNA, for example under induced conditions, or alternatively, a gene whose gene product can make up at least 0.2% (w/w) of the total cellular protein, or, in case of a secreted gene product, can be secreted to a level of at least 0.05 g/l.

also be used, for example the Moloney murine leukaemia virus long terminal repeat (MMLV LTR), the rous sarcoma virus (RSV) LTR promoter, the SV40 promoter, the human cytomegalovirus (CMV) IE promoter, herpes simplex virus promoters or adenovirus promoters.

Suitable yeast promoters include the *S. cerevisiae* GAL4 and ADH promoters and the *S. pombe* nmt1 and adh promoter. Mammalian promoters include the metallothionein promoter which can be induced in response to heavy metals such as cadmium. Viral promoters such as the SV40 large T antigen promoter or adenovirus promoters may also be used. All these promoters are readily available in the art.

Mammalian promoters, such as ß-actin promoters, may be used. Tissue-specific promoters, in particular endothelial or neuronal cell specific promoters (for example the DDAHI and DDAHII promoters), are especially preferred. Viral promoters may also be used, for example the Moloney murine leukaemia virus long terminal repeat (MMLV LTR), the rous sarcoma virus (RSV) LTR promoter, the SV40 promoter, the human cytomegalovirus (CMV) IE promoter, adenovirus, HSV promoters (such as the HSV IE promoters), or HPV promoters, particularly the HPV upstream regulatory region (URR). Viral promoters are readily available in the art.

A variety of promoters can be used that are capable of directing transcription in the host cells of the invention. Preferably the promoter sequence is derived from a highly expressed gene as previously defined. Examples of preferred highly expressed genes from which promoters are preferably derived and/or which are comprised in preferred predetermined target loci for integration of expression constructs, include but are not limited to genes encoding glycolytic enzymes such as triose-phosphate isomerases (TPI), glyceraldehyde-phosphate dehydrogenases (GAPDH), phosphoglycerate kinases (PGK), pyruvate kinases (PYK), alcohol dehydrogenases (ADH), as well as genes encoding amylases, glucoamylases, proteases, xylanases, cellobiohydrolases, ß-galactosidases, alcohol (methanol) oxidases, elongation factors and ribosomal proteins. Specific examples of suitable highly expressed genes include e.g. the LAC4 gene from *Kluyveromyces* sp., the methanol oxidase genes (AOX and MOX) from *Hansenula* and *Pichia*, respectively, the glucoamylase (glaA) genes from *A.niger* and *A.awamori*, the *A.oryzae* TAKA-amylase gene, the *A.nidulans* gpdA gene and the *T.reesei* cellobiohydrolase genes.

Examples of strong constitutive and/or inducible promoters which are preferred for use in fungal expression hosts are those which are obtainable from the fungal genes

expression vector as described above under conditions suitable for expression by the vector of a coding sequence encoding the polypeptide, and recovering the expressed polypeptide. Polynucleotides of the invention can be incorporated into a recombinant replicable vector, such as an expression vector. The vector may be used to replicate the nucleic acid in a compatible host cell. Thus in a further embodiment, the invention provides a method of making a polynucleotide of the invention by introducing a polynucleotide of the invention into a replicable vector, introducing the vector into a compatible host cell, and growing the host cell under conditions which bring about the replication of the vector. The vector may be recovered from the host cell. Suitable host cells include bacteria such as *E. coli*, yeast, mammalian cell lines and other eukaryotic cell lines, for example insect cells such as Sf9 cells and (e.g. filamentous) fungal cells.

Preferably the polypeptide is produced as a secreted protein in which case the DNA sequence encoding a mature form of the polypeptide in the expression construct is operably linked to a DNA sequence encoding a signal sequence. In the case where the gene encoding the secreted protein has in the wild type strain a signal sequence preferably the signal sequence used will be native (homologous) to the DNA sequence encoding the polypeptide. Alternatively the signal sequence is foreign (heterologous) to the DNA sequence encoding the polypeptide, in which case the signal sequence is preferably endogenous to the host cell in which the DNA sequence is expressed. Examples of suitable signal sequences for yeast host cells are the signal sequences derived from yeast MFalpha genes. Similarly, a suitable signal sequence for filamentous fungal host cells is e.g. a signal sequence derived from a filamentous fungal amyloglucosidase (AG) gene, e.g. the *A.niger* glaA gene. This signal sequence may be used in combination with the amyloglucosidase (also called (gluco)amylase) promoter itself, as well as in combination with other promoters. Hybrid signal sequences may also be used within the context of the present invention.

Preferred heterologous secretion leader sequences are those originating from the fungal amyloglucosidase (AG) gene (glaA - both 18 and 24 amino acid versions e.g. from *Aspergillus*), the MFalpha gene (yeasts e.g. *Saccharomyces* and *Kluyveromyces*) or the alpha-amylase gene (*Bacillus*).

The vectors may be transformed or transfected into a suitable host cell as described above to provide for expression of a polypeptide of the invention. This process may comprise culturing a host cell transformed with an expression vector as

selected from the group consisting of the genera *Aspergillus, Trichoderma, Fusarium, Disporotrichum, Penicillium, Acremonium, Neurospora, Thermoascus, Myceliophtora, Sporotrichum, Thielavia*, and *Talaromyces*. More preferably a filamentous fungal host cell is of the species *Aspergillus oyzae, Aspergillus sojae* or *Aspergillus nidulans* or is of a species from the *Aspergillus niger* Group (as defined by Raper and Fennell, The Genus *Aspergillus*, The Williams & Wilkins Company, Baltimore, pp 293-344, 1965). These include but are not limited to *Aspergillus niger, Aspergillus awamori, Aspergillus tubigensis, Aspergillus aculeatus, Aspergillus foetidus, Aspergillus nidulans, Aspergillus japonicus, Aspergillus oryzae* and *Aspergillus ficuum*, and also those of the species *Trichoderma reesei, Fusarium graminearum, Penicillium chrysogenum, Acremonium alabamense, Neurospora crassa, Myceliophtora thermophilum, Sporotrichum cellulophilum, Disporotrichum dimorphosporum* and *Thielavia terrestris*.

Examples of preferred expression hosts within the scope of the present invention are fungi such as *Aspergillus* species (in particular those described in EP-A-184,438 and EP-A-284,603) and *Trichoderma* species; bacteria such as *Bacillus* species (in particular those described in EP-A-134,048 and EP-A-253,455), especially *Bacillus subtilis, Bacillus licheniformis, Bacillus amyloliquefaciens, Pseudomonas* species; and yeasts such as *Kluyveromyces* species (in particular those described in EP-A-096,430 such as *Kluyveromyces lactis* and in EP-A-301,670) and *Saccharomyces* species, such as *Saccharomyces cerevisiae*.

Host cells according to the invention include plant cells, and the invention therefore extends to transgenic organisms, such as plants and parts thereof, which contain one or more cells of the invention. The cells may heterologously express the polypeptide of the invention or may heterologously contain one or more of the polynucleotides of the invention. The transgenic (or genetically modified) plant may therefore have inserted (typically stably) into its genome a sequence encoding the polypeptides of the invention. The transformation of plant cells can be performed using known techniques, for example using a Ti or a Ri plasmid from *Agrobacterium tumefaciens*. The plasmid (or vector) may thus contain sequences necessary to infect a plant, and derivatives of the Ti and/or Ri plasmids may be employed.

The host cell may overexpress the polypeptide, and techniques for engineering over-expression are well known and can be used in the present invention. The host may thus have two or more copies of the polynucleotide.

(e.g. phosphate, magnesium, potassium, zinc, iron, etc.). Optionally, an inducer (dependent on the expression construct used) may be included or subsequently be added.

The selection of the appropriate medium may be based on the choice of expression host and/or based on the regulatory requirements of the expression construct. Suitable media are well-known to those skilled in the art. The medium may, if desired, contain additional components favoring the transformed expression hosts over other potentially contaminating microorganisms.

The fermentation may be performed over a period of from 0.5-30 days. Fermentation may be a batch, continuous or fed-batch process, at a suitable temperature in the range of between 0°C and 45°C and, for example, at a pH from 2 to 10. Preferred fermentation conditions include a temperature in the range of between 20°C and 37°C and/or a pH between 3 and 9. The appropriate conditions are usually selected based on the choice of the expression host and the protein to be expressed.

After fermentation, if necessary, the cells can be removed from the fermentation broth by means of centrifugation or filtration. After fermentation has stopped or after removal of the cells, the polypeptide of the invention may then be recovered and, if desired, purified and isolated by conventional means. The proline specific endoprotease of the invention can be purified from fungal mycelium or from the culture broth into which the proline specific endoprotease is released by the cultured fungal cells.

In a preferred embodiment the polypeptide is obtained from a fungus, more preferably from an *Aspergillus*, most preferably from *Aspergillus niger*.

Modifications

Polypeptides of the invention may be chemically modified, e.g. post-translationally modified. For example, they may be glycosylated (one or more times) or comprise modified amino acid residues. They may also be modified by the addition of histidine residues to assist their purification or by the addition of a signal sequence to promote secretion from the cell. The polypeptide may have amino- or carboxyl-terminal extensions, such as an amino-terminal methionine residue, a small linker peptide of up to about 20-25 residues, or a small extension that facilitates purification, such as a poly-histidine tract, an antigenic epitope or a binding domain.

A polypeptide of the invention may be labelled with a revealing label. The revealing label may be any suitable label which allows the polypeptide to be detected.

truncation, and tyrosine, serine or threonine phosphorylation) as may be needed to confer optimal biological activity on recombinant expression products of the invention.

Preparations

Polypeptides of the invention may be in an isolated form. It will be understood that the polypeptide may be mixed with carriers or diluents which will not interfere with the intended purpose of the polypeptide and still be regarded as isolated. A polypeptide of the invention may also be in a substantially purified form, in which case it will generally comprise the polypeptide in a preparation in which more than 70%, e.g. more than 80%, 90%, 95%, 98% or 99% of the proteins in the preparation is a polypeptide of the invention.

Polypeptides of the invention may be provided in a form such that they are outside their natural cellular environment. Thus, they may be substantially isolated or purified, as discussed above, or in a cell in which they do not occur in nature, for example a cell of other fungal species, animals, plants or bacteria.

Removal or reduction of proline specific endoprotease activity

The present invention also relates to methods for producing a mutant cell of a parent cell, which comprises disrupting or deleting the endogenous nucleic acid sequence encoding the polypeptide or a control sequence thereof, which results in the mutant cell producing less of the polypeptide than the parent cell.

The construction of strains which have reduced proline specific endoprotease activity may be conveniently accomplished by modification or inactivation of a nucleic acid sequence necessary for expression of the proline specific endoprotease in the cell. The nucleic acid sequence to be modified or inactivated may be, for example, a nucleic acid sequence encoding the polypeptide or a part thereof essential for exhibiting proline specific endoprotease activity, or the nucleic acid sequence may have a regulatory function required for the expression of the polypeptide from the coding sequence of the nucleic acid sequence. An example of such a regulatory or control sequence may be a promoter sequence or a functional part thereof, i.e., a part which is sufficient for affecting expression of the polypeptide. Other control sequences for possible modification include, but are not limited to, a leader sequence, a polyadenylation sequence, a propeptide sequence, a signal sequence, and a termination sequence.

Modification or inactivation of the nucleic acid sequence may be performed by subjecting the cell to mutagenesis and selecting cells in which the proline specific

Alternatively, modification or inactivation of the nucleic acid sequence encoding a polypeptide of the present invention may be achieved by established anti-sense techniques using a nucleotide sequence complementary to the polypeptide encoding sequence. More specifically, production of the polypeptide by a cell may be reduced or eliminated by introducing a nucleotide sequence complementary to the nucleic acid sequence encoding the polypeptide. The antisense polynucleotide will then typically be transcribed in the cell and will be capable of hybridizing to the mRNA encoding the proline specific endoprotease. Under conditions allowing the complementary antisense nucleotide sequence to hybridize to the mRNA, the amount of the proline specific endoprotease produced in the cell will be reduced or eliminated.

It is preferred that the cell to be modified in accordance with the methods of the present invention is of microbial origin, for example, a fungal strain which is suitable for the production of desired protein products, either homologous or heterologous to the cell.

The present invention further relates to a mutant cell of a parent cell which comprises a disruption or deletion of the endogenous nucleic acid sequence encoding the polypeptide or a control sequence thereof, which results in the mutant cell producing less of the polypeptide than the parent cell.

The polypeptide-deficient mutant cells so created are particularly useful as host cells for the expression of homologous and/or heterologous polypeptides. Therefore, the present invention further relates to methods for producing a homologous or heterologous polypeptide comprising (a) culturing the mutant cell under conditions conducive for production of the polypeptide; and (b) recovering the polypeptide. In the present context, the term "heterologous polypeptides" is defined herein as polypeptides which are not native to the host cell, a native protein in which modifications have been made to alter the native sequence, or a native protein whose expression is quantitatively altered as a result of a manipulation of the host cell by recombinant DNA techniques.

In a still further aspect, the present invention provides a method for producing a protein product essentially free of proline specific endoprotease activity by fermentation of a cell which produces both an proline specific endoprotease polypeptide of the present invention as well as the protein product of interest. The method comprises adding an effective amount of an agent capable of inhibiting proline specific endoprotease activity to the fermentation broth either during or after the fermentation has been completed, recovering the product of interest from the fermentation broth, and optionally subjecting the recovered product to further purification. Alternatively, after

mixtures of casein and whey protein. Such mixtures of casein and whey protein may be used, for example, in ratios similar to those found in human milk. Furthermore, collagen based animal proteins forms a substrate because of the possibility to degrade these proteins to smaller molecules hereby debittering animal meat extracts or improving the uptake of proline and hydroxyproline residues with benefits on the joints of athletes.

The enzyme mixture according to the invention may also be used to hydrolyze proteinaceous materials of plant origin such as, for example, wheat gluten malted or unmalted barley or other cereals used for making beer, soy milk, concentrates or isolates thereof, maize protein concentrates and isolates thereof, and rice proteins.

The invention will be further illustrated by the following non-limiting Examples.

Examples

Materials and Methods

Beta-casein from bovine milk (lyophilised,essentially salt-free powder) with a minimum 90% beta-casein was obtained from Sigma. Collagen (Type 1, insoluble from bovine achilles tendon) was also obtained from Sigma.

Sodium caseinate (Miprodan 30®)) was obtained from MD Foods (Viby,Denmark). Sweet whey concentrate, nonpasteurised, 10% ds, 35% protein was obtained from Borculo Domo (Zwolle, The Netherlands)

A low bitterness whey hydrolysate Vitalarmor® 800 LB as well as whey protein enriched in beta-lactoglobulin (Protarmor® 905) was obtained from Armor Proteines (Saint-Brice-en -Cogles,France). Other commercial hydrolysates were obtained from the producer or purchased in pharmacies.

Soy isolate was obtained as Soyamin® 90 HV from Lucas Meyer, Hamburg, Germany.

Subtilisin from *B.licheniformis* (Delvolase®, 560 000 DU per gram) was obtained from DSM Food Specialities (Seclin, France). Sumizyme® LP 75.000 was obtained from Shin Nihon (Anjyo, Japan). Flavourzyme® 1000L was obtained from NOVO Industries, Bagsvaerd, Denmark. Thermolysin (Thermoase; a heat stable metallo-endoprotease from Bacillus thermoproteolyticus Rokko with an activity of 14000 PU/ mg as produced by Daiwa Kasei, Osaka, Japan)

Proline-specific endoprotease from *Flavobacterium meningosepticum* and cloned in *E.coli* was isolated using known plasmid constructs and enzyme purification methods (T.Diefenthal and H.Dargatz, World Journal of Microbiology &Biotechnology 11, 209-212 (1995)) The enzymatic activity was tested on CBZ-Gly-Pro-pNA 0.26 mM in phosphate

After incubation the IPG strip was trimmed to fit the gel type mentioned and ran with 10x diluted TGS buffer (BioRad). After the run the gel was stained with Sypro Ruby (Molecular Probes,Leiden, The Netherlands) for 3-4 hours and washed with Milli Q water for 2 hours. Imaging was performed on The Imager (Appligene). The largest spot was cut out, washed several times with 50 millimoles/litre ammonium bicarbonate, incubated overnight at 37 degrees C with sequencing grade trypsin (nr.1047841, Boehringer Mannheim). Peptides were extracted from the gel piece by washing several times with acetonitrile/water containing formic acid (50/50/5, v/v/v). The samples were dried using a vacuumcentrifuge (New Brunswick Scientific, The Netherlands) and stored at –20° C , until analysis.

LC/MS Analysis

HPLC (high performance liquid chromatography) using a Qtof-2 (Micromass, Manchester, UK) mass spectrometer was used to separate the peptides formed during digestion with trypsin. 5 microliter of the peptide solution was trapped on a micro-precolumn, C18, 5*0.3 mm (MCA30-05-C18, LC Packings, Amsterdam, Netherlands) using Milli Q water containing 0.1 % of formic acid at a flow-rate of 20 microliter/min. The peptides were then eluted from the precolumn, using a fast gradient of 0.1% formic acid in Milli Q water (Millipore, Bedford, MA, USA; Solution A) and 0.1% formic acid in acetonitrile (Solution B). The gradient started at 100% of Solution A and increased to 60% of solution B in 20 minutes and was kept at the latter ratio for another 5 minutes. The flow rate used during elution of the peptides was 200 nl/min. Using LC/MS/MS analysis partial amino acid sequences of the *A. niger* proline-specific endopeptidase could be determined, by *de novo* sequencing of suitable peptides.

HPLC using an ion trap mass spectrometer (Thermoquest™, Breda, the Netherlands) coupled to a P4000 pump (Thermoquest™, Breda, the Netherlands) was used in characterising the enzymatic protein hydrolysates produced by the inventive enzyme mixture. The peptides formed were separated using a PEPMAP C18 300A (MIC-15-03-C18-PM, LC Packings, Amsterdam, The Netherlands) column in combination with a gradient of 0.1% formic acid in Milli Q water (Millipore, Bedford, MA, USA; Solution A) and 0.1% formic acid in acetonitrile (Solution B) for elution. The gradient started at 100% of Solution A and increased to 70% of solution B in 45 minutes and was kept at the latter ratio for another 5 minutes. The injectionvolume used was 50 microliters, the flow rate was 50 microliter per minute and the column temperature was

masses between 400 and 2000 daltons were further characterised by the MS/MS analysis. However, the databank used to obtain peptide sequence information on whey or casein derived hydrolysates consisted of cow milk protein sequences only.

Determination of the molar fraction of peptides (%) carrying a carboxyterminal proline.

LC/MS/MS can be used for the analysis of the C-terminus of a peptide. With an algorithm in which the peptide's molecular mass (analyzed with LC/MS) and its (partial) amino acid sequence (analyzed with LC/MS/MS) are linked with automatic search procedures within protein databanks, complex peptide mixtures can be analyzed. These options have enabled us to quantify the incidence of peptides carrying a carboxy terminal proline residue. Owing to the limitations set by the PEPMAP peptide separation column used, only peptides with a molecular weight between roughly 400 and 2000 Dalton are analysed using this technique. Fortunately, in protein hydrolysates the majority of the peptides have such molecular weights.

To determine in a protein hydrolysate the molar fraction of peptides carrying a carboxyterminal proline, individual peptide peaks eluting from the PEPMAP column are selected and partial carboxyterminal amino acid sequences are determined using the techniques specified above. Analysis of at least 20, preferably at least 30 and more preferably between 40 to 60, for example 50 of the most abundant, randomly choosen peptides thus provides insight in the frequency in which peptides carrying a proline residue at the carboxyterminus of the peptide occur. The quotient of the number of peptides found to carry a carboxyterminal proline residue times 100 and the total number of peptides analysed thus provides the molar fraction of peptides (%) carrying a carboxyterminal proline.

Determination of the molar fraction (%) of proline in the protein substrate used to generate the hydrolysate.

Fatty material as can occur in infant formulae products was first removed by hexane extraction as detailed in the paragraph describing LC/MS analysis of infant formulae and commercial protein hydrolysates. Acid hydrolysis of the protein substrate to convert the proteins present into free amino acids, was achieved by making a suspension of 100 milligrams of proteinaceous material in 2 milliliters 6 N HCl. Acid hydrolysis was carried out for 22 hours at 112 degrees C in an oxygen free atmosphere. After centrifugation the supernatant was diluted 10 times in dilute HCl. After this hydrolysis the

available in an almost pure form. As such, beta-casein offers an excellent test substrate for studying the relationship between enzyme cleavage sites and the length of various peptides formed during enzyme hydrolysis.

This Example demonstrates that despite the broad spectrum character of subtilisin, the addition of a very specific enzyme like a proline-specific endoprotease can have a major impact on the size of the beta-casein fragments formed. Improved yields for casein fractions upon incubation with subtilisin in combination with a proline-specific endoprotease can therefore be obtained. Beta-casein is relatively rich in proline as acid hydrolysis followed by amino acid analysis carried out according to the Materials & Methods section revealed that its molar fraction of proline is 14% (moles of proline/moles of all amino acids as specified in the Materials&Methods section).

Beta-casein powder (Sigma) was dissolved at a concentration of 10% (w/w) together with 0.1% (w/w) Delvolase™ in a 0.1 mol/liter phosphate buffer pH7.0. After an incubation of 24 hours at 45°C in a shaking waterbath, the reaction was stopped by heating the solution for 15 minutes at 90°C. To one half of the solution (1ml containing 100milligrams of beta-casein) 100 microliter of proline-specific endoprotease from *F. meningosepticum* (corresponding to 4 units according to the procedure described in World Journal of Microbiology&Biotechnology, Vol 11, pp209-212) was added and the reaction was continued for another 24 hours at 45°C. After another heat shock at 90°C, samples of both the Delvolase™ and the Delvolase™ + proline-specific endoprotease treated beta-casein material were analysed by LC/MS equipment as specified in the Materials and Methods section.

In the sample digested with Delvolase alone, the LC/MS/MS analysis identified 40 peptides covering various parts of the beta-casein molecule. Together these peptides accounted for 79% of the total beta-casein sequence. Different retention times of the peptides on the C18 column could be traced back to peptide lengths ranging from 2 to 23 amino acid residues. Glutamine proved to be the most frequently occurring carboxy-terminal residue (10 out of 40 peptides). None of the peptides analysed could be shown to have proline as the carboxy terminal residue.

By contrast, the sample digested with Delvolase™ and proline-specific endoprotease generated 28 identifiable peptides from beta-casein. Together these peptides covered 63% of the total beta-casein protein sequence. Peptide size distribution was remarkably homogeneous, as the peptides ranged in length only between 3 and 9 residues. Within this peptide population, glutamine was the carboxy-

gram of KCl, 0.5 gram of MgSO4. 7H2O, 0.01 gram of FeSO4. 7H2O, 5grams of glucose, 15grams of collagen (Sigma) and distilled water added to obtain a volume of 1 litre. The inoculum for each experiment was prepared by a method in which the spores of fungi growing on an agar slant (5 days old) were taken up in 5 milliliters of sterile water. Of the latter suspension, 2%(v/v) was used for inoculation of the pH 6.5 medium. Growth was allowed for 100 hours at 28 degrees C with shaking after which the culture was filtrated and samples of the clear filtrate were incubated with the synthetic peptide Z-Ala-Pro-pNA (Bachem; Bubendorf, Switzerland) at pH 5.0, 50 degrees C. Samples capable of releasing pNA were identified by measuring the increase in absorbance at 410 nanometer. Positive strains yielding relatively high activities were further investigated.

Strain G-306 excreted a proline-specific endoprotease and was identified as *Aspergillus niger* Van Tieghem var. *niger*. This particular strain was used for isolation, purification and further characterisation of a proline-specific endoprotease. To purify the enzyme 1 liter of culture supernatant was applied to a 400 milliliter bacitracin-silochrome column equilibrated with 0.05 mol/litre sodium acetate pH 5.0. Proteases bound to the column were eluted using the acetate buffer supplemented with 1 mol/litre of NaCl and 10% (v/v) isopropanol (J.Appl.Biochem.,1983 pp420-428). Active fractions were collected and dialysed against distilled water and applied on a 200 milliliter bacitracin-Sepharose column, again equilibrated with acetate buffer. As before, elution was carried out using the acetate buffer supplemented with NaCl and isopropanol. Active fractions were collected, dialysed against a 5 millimol/litre acetate buffer pH 5.0 and then concentrated by means of ultrafiltration with a Amicon PM-10 membrane. To obtain an almost completely pure proline-specific endoprotease, the concentrated liquid was chromatographed over a Superdex™ 75 column eqiulibrated with the 0.05 mol/litre sodium acetate buffer pH 5.0 and supplemented with 0.5 mol/litre NaCl.

Further experiments carried out with the purified enzyme indicated a molecular weight around 66.6 kDalton, an IEP around pH 4.2, a pH optimum around 5.0 and an almost 100% thermostablity upon incubation for 4hours at 50 degrees C.

To obtain partial amino acid sequences of the enzyme, the enzyme preparation isolated was first subjected to two-dimensional gel electrophoresis according to the procedure described in the Materials &Methods section. The largest spot was cut out, incubated with trypsin and eluted. The recovered peptides were then subjected to

Protease : in the order of 650 PU; carboxypeptidase : in the order of 0.01 unit and proline-specific endopeptidase : in the order of 0.03 milli-units.

Because the original *Aspergillus oryzae* FS 1-32 preparation was not available, two commercial enzyme preparations, also derived from *Aspergillus oryzae*, were used in the present Example. Moreover, a chromatographically purified proline-specific endoprotease isolated from *Aspergillus niger* (see Example 3) was used to achieve an overdosing of the acid proline-specific endoprotease.

The enzymatic activities of the various preparations were measured according to the procedures provided in JP5015314 and are provided below.

- Sumizyme LP 75.000, a commercial *Aspergillus oryzae* enzyme preparation known to be rich in endoproteolytic activity.
  Enzymatic activities as assessed according to the methods of JP5015314:
  Protease : 226 PU/gram product ; carboxypeptidase : 21 units/gram product ; prolyl-endopeptidase : 430 milli-units/gram product
- Flavourzyme 1000L , a commercial *Aspergillus oryzae* enzyme preparation known to be rich in exoproteolytic activity.
  Enzymatic activities as assessed according to the methods of JP5015314:
  Protease : 332 PU/gram product ; carboxypeptidase : 10 units/gram product ; prolyl-endopeptidase : not detectable
- Chromatographically pure proline-specific endoprotease obtained from *Aspergillus niger* and isolated as described in Example 3.
  Enzymatic activities as assessed according to the methods of JP5015314:
  Protease : not detectable ; carboxypeptidase : not detectable ; prolyl-endopeptidase : 45 milli-units/milliliter.

From these data it is evident that although Sumizyme and Flavourzyme are well known for their high proteolytic activities, none of them can provide the same very high ratio of (endo)protease to carboxypeptidase activity as quoted in JP5015314. Surprisingly Sumizyme LP 75.000 was found to contain a considerably higher activity of proline-specific endoprotease than the one reported in JP5015314.

The various enzyme preparations were incubated according to the protocol described in JP5015314 but standardised according to the desired carboxypeptidase activity (0.01 unit per gram substrate). Soy isolate (Soyamin 90 HV) was used as the substrate in

is about 50 times higher than the activity recorded with strain FS 1-32 but also yields a molar fraction of approx 10% of soy peptides carrying a carboxy terminal proline. These data were confirmed by analysing the number of proline residues which are present in the peptides but not in the carboxy terminal position. Flavourzyme contains no detectable proline-specific endpoprotease but yields among the peptides generated and suitable for analysis with the LC/MS technique a molar fraction of 6% of peptides carrying a proline at the carboxy terminal end. If combined with a proline content of approx. 5 % of this soy protein isolate, these three observations indicate that the presence and the activity of the proline-specific endoprotease in combination with the carboxypeptidase activity has a minor effect on the molar incidence of carboxy terminal proline residues only. So, it is hard to imagine that the debittering effect described in JP5015314 and ascribed to a proline-specific endoprotease activity of 0.03 milli-units only can be linked to a high incidence of peptides carrying proline as the carboxy terminal amino acid residue.

Example 5

Increased dosages of proline-specific endoprotease and its effects on the hydrolysis of soy protein.

In this Example it is demonstrated that high levels of a proline-specific endoprotease are required to generate soy hydrolysates containing a significant amount of peptides carrying a carboxy terminal proline residue.The overall design of these experiments was identical to the ones described in Example 4. Again soy protein isolate was incubated with Sumizyme LP 75.000 standardised according to the desired carboxypeptidase activity of 0.01 unit per gram soy protein and under conditions described in JP5015314. The incubation took place for either 2.5 or 5.0 hours at pH 5 and 50 degrees C and was stopped by keeping the material for 10 minutes at *100* degrees C. Subsequently some of the material incubated for 5 hours was obtained and its pH was increased to 7.0. From this material 3 samples were obtained to which different portions of the *E .coli* produced *F. meningosepticum* proline-specific endoprotease were added . To the first sample 1.5 milli-units of proline-specific endoprotease (according to JP5015314 but measured at pH 7.0 and 30 degrees C to accomodate the pH and temperature optimum of the *E.coli* derived proline-specific endoprotease) were added, to the second sample 150 milli-units were added and to the third sample 15 000 milli-units were added and then the samples were again incubated

Example 6

Molar incidence of peptides carrying proline as the carboxy terminal residue in commercial hydrolysates.

As described earlier, LC/MS/MS can be used for the analysis of the C-terminus of a peptide. With an algorithm in which the peptide's molecular mass (analyzed with LC/MS) and its (partial) amino acid sequence (analyzed with LC/MS/MS) are linked with automatic search procedures within protein databanks, complex peptide mixtures can be analyzed.

In this Example these possibilities were used to analyse a number of commercial infant formulae products as well as commercial protein hydrolysates for the molar incidence of peptides carrying carboxy terminal proline residues which have a molecular weight between 400 and 2000 daltons.

The following products were analysed.

1. Nidal® HA 1(Nestle), containing 11.5 g whey-protein hydrolysates per 100 g powder
2. Alfare®(Nestle), containing 16.5 g whey-protein per 100 g powder
3. Nutrilon® Pepti Plus(Nutricia), containing 13.5 g whey-protein per 100 g powder
4. Nutrilon® Pepti Junior(Nutricia), containing 16.5 g whey-protein hydrolysates per 100 g powder
5. Aptamil® HA(Milupa), containing 12.3 g whey-protein and casein hydrolysates per 100 g powder
6. Pregomin®(Milupa), containing 13.3 g of probably soy and collagen hydrolysates per 100 g powder
7. Nutramigen®(Mead Johnson), containing 14.0 g of probably casein hydrolysates 100 g powder
8. Vitalarmor® 800 LB(Armor Proteines), containing 100% whey-protein hydrolysates
9. WPH 916(New Zealand Milk Products), containing 100% whey-protein hydrolysates
10. WE80 BG(DMV International), containing 100% whey-protein hydrolysates

As the infant formulae contain approx.15% of protein hydrolysate plus fats (25%) and carbohydrates (50%), a hexane extraction of these products to remove the fat phase proved to be indispensible. The pure hydrolysates could be used as such.

To link the partial protein sequences obtained with sequences of known proteins, a databank containing cow milk protein sequences only was used for all samples except the Pregomin sample. The Pregomin sample was analysed using a databank containing

Table 3: Molar incidence of peptides carrying carboxy terminal proline in commercial products.

Infant formulae	Protein basis	Number of peptides analysed	Molar fraction of peptides carrying C-terminal proline (%)	Molar fraction of proline in protein basis (%)
Nidal HA 1	Whey	49	0	5
Alfare	Whey	50	2	7
N. Pepti Plus	Whey	74	4	7
N. Pepti Junior	Whey	72	3	7
Aptamil HA	Whey / casein	69	3	9
Pregomin	Soy / collagen	41	7	8
Nutramigen	Casein	32	22	11
Pure hydrolysates				
Vitalarmor 800 LB	Whey	54	6	6
WPH 916	Whey	69	0	5
WE 80 BG	Whey	76	3	8

From the data presented in Table 3 it is clear that in the popular whey hydrolysates the molar incidence of peptides carrying carboxy terminal proline residues is low. If we also take the proline content of whey into account, we conclude that none of the commercial whey based products contains a molar fraction of peptides carrying carboxy terminal proline residues which is higher than the molar fraction of proline occurring in the protein basis. Typically the molar fraction of peptides carrying a carboxy terminal proline in these whey based commercial products is 5% or lower.

Looking at the molar incidence of carboxy terminal proline residues in a casein based product like Nutramigen, we see a substantial higher level than can be found in the whey based products even if the relatively high proline content of casein is taken into account. However, comparing the Nutramigen product on the one hand with the beta-casein hydrolysate made by incubation with subtilisin and a proline-specific endo-

Table 4: Enzyme dosage and molar incidence of peptides carrying carboxyterminal proline.

Temperature	Dosage of proline-specific endoprotease	Number of peptides analysed	Number of peptides carrying proline in C-terminal position	Molar incidence of peptides with proline at C-terminal position (%)
30°C	0 U/g	40	2	4
	87 U/g	33	12	52
	170 U/g	46	19	53
45°C	0 U/g	45	0	0
	87 U/g	49	15	36
	170 U/g	29	13	50

From this Table, it appears that at 45°C the molar incidence of peptides carrying proline at their C-terminus increases with the dose of the proline-specific endoprotease. Using the highest enzyme dosages, up to 50 % of the peptides obtained from this whey product could be shown to carry a carboxy terminal proline residue. When the incubation is performed at 30°C, the molar incidence of peptides carrying a carboxy terminal proline residue can reach 52% with 87 units/gram substrate and is hardly increased with higher doses of the enzyme. The higher incidence reached with 87U/g at 30°C compared to 45°C might be explained by a low thermostability of the *E.coli* enzyme.

Example 8

Taste and composition of whey hydrolysates produced with and without proline-specific endoprotease.

In this Example a proline-specific endoprotease obtained from *E.coli* was used in combination with subtilisin (Delvolase) to produce a whey hydrolysate of low bitterness. Using the data generated in Example 7 the dosage of the proline-specific endoprotease was chosen such that only a marginal increase of peptides carrying a

Of the 24 persons participating in the first test, 17 rated the sample obtained with the combination of subtilisin and proline-specific endoprotease as less bitter than the sample obtained with subtilisin alone. This result is statistically significant and confirms the debittering activity of a proline-specific endoprotease, even if applied at relatively low concentrations (cf Example 7). Worthwile to note is that these "low"enzyme concentrations are several orders of magnitude higher than the enzyme dosages applied in patent JP5015314 and for which a debittering effect was claimed.

In the second paired sample comparison, 19 out of the 24 participants rated the sample treated with the combination of subtilisin and proline-specific endoprotease as less bitter than the commercial Vitalarmor 800LB product. The latter observation is statistically also significant and illustrates the economical value of the hydrolysates and enzyme mixtures of the invention.

The hydrolysates obtained with or without the proline-specific endoprotease were analysed by LC/MS as described before. In the hydrolysate obtained with the subtilisin alone, 41 peptides were analysed. It came out that none of these peptides carried a carboxy terminal proline residue despite the fact that 18 peptides were shown to contain at least one proline residue.

In the hydrolysate obtained with the combination of subtilisin and proline-specific endoprotease 31 peptides were analysed and 6 were shown to carry a carboxy terminal proline residue. This observation, which is in line with what could be expected on the basis of the results obtained in Example 6, shows that as the result of the incubation with the proline-specific endoprotease the molar incidence of peptides bearing a carboxy terminal proline residue was increased from 0 to 19%. As the sensory analysis of the latter products has demonstrated a statistically significant reduced bitterness, this experiment clearly links a slight increase in the molar incidence of carboxy terminal proline residues with reduced bitterness.

Apart from decreasing the level of bitterness, this incubation with a low level of proline-specific endoprotease could also be shown to decrease the peptide length of the hydrolysate. In the hydrolysate treated with Delvolase alone, the LC/MS analysis revealed that peptides vary in length from 4 to 14 amino-acids with an average length of 7.5 amino-acids. In the hydrolysate treated with the combination of Delvolase and the proline-specific endoprotease, the peptide length could be shown to vary from 4 to 12 amino-acids with an average length of 6.1 amino-acids. These reduced peptide lengths

PCR was decreased and several bands were amplified, although most of these bands were also detected in control PCR reactions lacking one of the primers. Several of these PCR products were cloned into the general cloning vector pCR2.1 (Invitrogen, Groningen, The Netherlands), and the DNA sequence of these fragments was determined. Unfortunately none of the cloned fragments coded for the gene encoding proline specific endoprotease.

Additionally, many other adjustments to the PCR protocol were made such as the use of a different polymerase, increasing primer- or template-concentration, a touch-down PCR and introduction of a hot start, but none of these protocols yielded a specific fragment of the gene encoding the proline specific endoprotease. To minimize the obvious risks of this uncertain approach, it was decided to try another, less well known cloning procedure.

3'-RACE

Since none of our attempts to amplify the gene encoding the proline specific endoprotease from *A. niger* G306 genomic DNA were successful, we decided to use a different approach in which RNA is used as the template for cDNA synthesis. The approach of cloning an unknown gene using 3'-RACE, 5'-RACE and amplification of the complete open reading frame, has been described in WO9938956. The advantage of this procedure, compared to the direct PCR procedure described above, is that an additional priming site is introduced at the 3'-end of the cDNA, so that only a single gene-specific oligonucleotide plus an universal primer is required to amplify part of the coding sequence, instead of two degenerate primers. Additionally, using cDNA as template circumvents problems in amplification due to introns. The use of cDNA as template in the amplification reaction also increases the concentration of the template compared to amplification from genomic DNA.

According to this approach, *A. niger* G306 was grown in a medium containing collagen as sole carbon source to induce the expression of the gene encoding for proline specific endoprotease. Medium composition is described in the Materials and Methods section. Young mycelium was harvested after 48 hr growth at 34 °C, and used for the isolation of total RNA. To this end, mycelium was harvested using filtration through Miracloth filtration wrap and washed with ice cold sterile demiwater. Mycelium (250 mg) was frozen immediately in liquid nitrogen and ground to a fine white powder using mortar and pestle. The white powder was transferred to a sterile 15 ml Greiner tube and total

deduced protein sequence of 526 amino acids is depicted in SEQ_ID 2. Peptide ATTGEAYFE appeared to be completely correct. Peptide DGAPEGTST is also correct but is encoded by genomic DNA that is interrupted by an intron (see SEQ_ID 15 and example 11 for the cloning and sequence of genomic DNA of Aspergillus niger CBS513.88). The other two peptides incorporate errors due to the LC/MS/MS approach which has been used for their characterization (see Example 3). Despite these uncertainties we successfully selected and identified the desired genetic information encoding the proline specific endoprotease from *Aspergillus* for the first time.

The novelty of the proline specific endoprotease from *Aspergillus* was confirmed by BLAST searches to well known databases such as SwissProt, PIR and trEMBL. No strong identity of this protein with any other protein can be detected when compared to the protein sequence databases.

Example 10

Overexpression of the gene encoding proline specific endoprotease, and isolation of the proline specific endoprotease

The entire open reading frame of the gene encoding proline specific endoprotease was PCR amplified from cDNA of *A. niger* G306 using the primers 5'-ATGCGTGCCTTCTCCGCTGTC-3' and the AUAP primer (Life Technologies). The obtained PCR fragment was cloned into the cloning vector pCR2.1 (Invitrogen). The resulting plasmid was digested with *Eco*RI and the fragment containing the endo-Pro gene was cloned into the *Eco*RI site of expression vector pGBFIN-11 (WO9932617). The resulting clones were checked by restriction with *Xho*I, which yields a fragment of ~0.65 kb when the fragment is inserted in the correct orientation. The resulting plasmid is shown in Figure 1 and was named pGBFIN11-EPO.

A. niger CBS 513.88 was used as host for the over-expression of the gene encoding the proline-specific endoprotease. Therefore, the expression vector pGBFIN11-EPO was linearized by digestion with *Not*I, which removes all *E. coli* derived sequences from the expression vector. The digested DNA was purified using phenol:chloroform:isoamylalcohol (24:23:1) extraction and precipitation with ethanol. The *A. niger* transformation procedure is extensively described in WO 98/46772. It is also described how to select for transformants on agar plates containing acetamide, and to select targeted multicopy integrants. Preferably, *A. niger* transformants containing

On the basis of the low nucleotide sequence homology between the *F.meningosepticum* and the *A. niger* gene encoding proline specific endoprotease, cross hybridization between these two genes can be excluded. To get an impression of the conservation of the *A. niger* specific nucleotide sequence in more related microorganisms, the following strains were selected for a hybridization experiment. The fungal species *Aspergillus niger* CBS102.12, *Aspergillus niger* CBS513.88, *Aspergillus niger* G306, *Aspergillus carbonarius* ATCC1025, *Aspergillus sojae* DSM2809, *Aspergillus ochraceus* ATCC18500, *Aspergilus acculeatis* CBS101.43, *Verticillium psalliotae* CBS396.58, *Phialophora mustea* CBS142.41, *Penicillium chrysogenum* URCM237, *Phoma exigua* CBS431.74, *Microsporum gallinae* CBS221.55, *Acremonium strictum* ATCC20371, *Rhizomucor miehei* CBS370.65, *Alternaria alternata* CBS103.33, *Talaromyces emersonii* CBS393.64, *Cladosporium chlorocephalum* CBS213.73, *Cladosporium tenuissinum* CBS117.79, and *Trichoderma reesii* ATCC26921 were cultivated in 100 ml PDB (Potato Dextrose Broth, Difco) at 30 °C (except for the Talaromyces strain which was grown at 50 °C) and shaken at 220 rpm.

When cultures were sufficiently grown, mycelial mass was harvested by filtration through Miracloth filter, washed with 10 mM KPi buffer (pH 7.0) and dried between filterpaper. Mycelium was ground under liquid nitrogen with a mortar and pestle, until a fine white powder was obtained. Subsequently, chromosomal DNA was isolated using the PureGene kit (Gentra Systems, Minneapolis USA) according to instructions by the supplier.

Saccharomyces cerevisiae ATCC20785 was used as negative control in the experiment and cultivated in YePD at 30 °C and shaken at 220 rpm.

For preparation of a Southern blot, chromosomal DNA of all species was digested with XhoI and restriction fragments were separated by agarose gelelectrophoresis on a 0.8% agarose gel in TAE buffer. After separation, DNA fragments were blotted to nitrocellulose (0.2 μm, Schleicher & Schuell) membranes by conventional procedure (Sambrook et al. (1982): Molecular cloning; a laboratory manual, ISBN 0-87969-309-6), and the blot was backed for 2 hours at 80 °C.

The probe for hybridization was synthesized with PCR on pGBFIN11-EPO as template using primers 5'-ATGCGTGCCTTCTCCGCTGTC-3' and the AUAP primer. About 30 nanograms of the cDNA fragment was labeled with ^{32}P-alpha-dATP (Amersham, England) with the RadPrime DNA labeling system (Life Technologies) according to the supliers instructions. After labeling unincorporated dNTP's were removed by purifying

specific endoprotease from other species and strains using hybridization to the cDNA of this gene from Aspergillus niger G306.

The deduced coding sequence and amino-acid sequence of the proline-specific endoprotease of CBS513.88 is depicted in SEQ_ID 16 and SEQ_ID 17 respectively.

Table 6: Heterologous hybridization of the A. niger endo-Pro gene to chromosomal DNA of various fungi.

Species	*Hybridization*
Aspergillus niger CBS102.12	+++
Aspergillus niger CBS513.88	+++
Aspergillus niger G306	+++
Aspergillus carbonarius ATCC1025	+++
Aspergillus sojae DSM2809	+
Aspergillus ochraceus ATCC18500	++
Aspergilus acculeatis CBS101.43	+
Saccharomyces cerevisiae ATCC20785	-
Verticillium psalliotae CBS396.58	-
Phialophora mustea CBS142.41	+
Penicillium chrysogenum URCM237	-
Phoma exigua CBS431.74	-
Microsporum gallinae CBS221.55	-
Acremonium strictum ATCC20371	-
Rhizomucor miehei CBS370.65	+
Alternaria alternata CBS103.33	+
Talaromyces emersonii CBS393.64	+
Cladosporium chlorocephalum CBS213.73	-
Cladosporium tenuissinum CBS117.79	-
Trichoderma reesii ATCC26921	+

Example 12.

Enzyme mixture obtained from *Aspergillus oryzae* FS 1-32 and its effects in the hydrolysis of soy protein.

Japanese patent JP5015314 discloses a crude enzyme preparation obtained from *Aspergillus oryzae* FS 1-32 containing major quantities of a non-specified endoproteolytic activity and minor quantities of a proline-specific endoprotease. This

The culture medium according to EP 0 522 428 contains the following components : acid casein (Armor Proteins, France) 25.4 grams/liter, roasted soybean flour (Cargill, Netherlands) 8.6 grams/liter, wheat bran (Zonnatura, Netherlands) 15.0 grams/liter, corn starch 20.0 grams/liter, tannic acid (Omnichem) 16.0 grams/liter and KH_2PO_4 26.6 grams/liter. Because the recommended tannic acid (to stimulate the formation of the proline-specific endoprotease) was not specified in EP 0522428, two kinds of tannic acids, i.e. BREWTAN C and TANAL W2 (both from Omnichem (Wetteren, Belgium) were used. Finally the pH value of the culture medium was adjusted with phosphoric acid (20%) to 4.5 and then divided in portions of 100 ml in 500 ml shake flasks with baffles. Flasks were sterilized for 30 minutes at 121°C.

After inoculation with 1 milliliter of the pre-grown inoculation medium, the cultures were incubated for 2 and 4 days at 32°C, 250 rpm. To remove the biomass, the culture broths were filtered over a Whatman glass microfibre filter (cat no 1820090) which were then stored at –20°C. Part of this frozen material was lyophilized and used for activity measurements as well as incubations with soy protein.

The activities of the prolyl-endopeptidase, carboxypeptidase and endoprotease in the lyophilized materials were measured exactly as described in JP5015314. The samples that had been fermented for 2 days showed appreciably higher enzyme activity levels then the samples that had been fermented for the recommended 4 days so that it was decided to use these 2 days samples for the final incubation with soy protein. Enzyme activity data of those samples showing the highest prolyl-endopeptidase activities are shown hereunder.

Table7. Enzyme activities per gram of lyophilized material

Sample		Prolyl-endopeptidase activity	Carboxypeptidase activity	Protease activity
		[mU/g]	[U/g]	[PU/g]
1	+Brewtan	2.87	4.99	609
3	+Tanal	2.38	3.68	595
4	+Tanal	6.30	7.79	592

The propyl-endopeptidase and the carboxypeptidase activities measured in the samples 1, 3 and 4 are comparable with the figures provided in JP5015314. However, the

Table 8: Soy protein treated with several enzymes.

Enzyme units per gram substrate	Number of peptides analysed	Molar fraction of peptides with proline at C-terminus (%)
None (reference)	73	3
Sample 1 (2.0 mg) Protease (PU): 1.20 Carboxypep (U):0.01 PEP(milli-Units): 0.006	76	1
Sample 3 (2.7mg) Protease (PU): 1.60 Carboxypep (U):0.01 PEP(milli-Units): 0.006	78	3
Sample 4 (1.3mg) Protease (PU): 0.80 Carboxypep (U):0.01 PEP(milli-units): 0.008	70	2
JP5015314 Protease: 650 Carboxypep: 0.01 PEP(milli-units):0.03		

PEP : prolyl-endopeptidase or proline-specific endoprotease.

From the data obtained it is obvious that the incubation of soy protein with the crude enzyme preparation obtained from *Aspergillus oryzae* FS 1-32 doesnot result in a significant increase of the molar fraction of peptides carrying a carboxyterminal proline residue. So the debittering effect described in JP5015314 cannot be attributed to a high incidence of such peptides in the final hydrolysate.

Applicant's or agent's file reference number 20001WO	International application No.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 59, line 11 .

B. **IDENTIFICATION OF DEPOSIT** Further deposits are identified on an additional sheet ☐

Name of depositary institution
CENTRAAL BUREAU VOOR SCHIMMELCULTURES (CBS)

Address of depositary institution *(including postal code and country)*
Oosterstraat 1
P.O.Box 273
3740 AG BAARN
The Netherlands

Date of deposit	Accession Number
10 september 2001	CBS 109712

C. **ADDITIONAL INDICATIONS** *(leave blank if not applicable)* This information is continued on an additional sheet ☐

We inform you that the availability of the microorganism identified above, referred to Rule 13bis PCT, shall be effected only by issue of a sample to an expert nominated by the requester until the publication of the mention of grant of the national patent or, where applicable, for twenty years from the date of filing if the application has been refused, withdrawn or deemed to be withdrawn.

D. **DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE** *(if the indications are not for all designated States)*

E. **SEPARATE FURNISHING OF INDICATIONS** *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☐ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

Form PCT/RO/134 (July 1992)

9. Use of a protein hydrolysate according to any one of claims 1 to 8 in a food or drink.

10. Use according to claim 9, wherein the food or drink comprises from 5% to 10% (w/v) of the protein hydrolysate.

11. Use according to claim 9 or 10 wherein the food has reduced bitterness and/or low antigenicity.

12. Use according to claim 11, wherein the food comprises an infant formula.

13. A method of enzymatically producing a protein hydrolysate from a protein substrate, wherein the protein substrate is incubated with a proline-specific endoprotease to produce a protein hydrolysate enriched in peptides having a carboxy terminal proline.

14. A method according to claim 13, wherein the protein substrate is incubated either sequentially or concomitantly with the proline-specific endoprotease and another endoprotease.

15. A method according to claim 13, wherein the protein substrate is incubated either sequentially or concomitantly with the proline-specific endoprotease and the other endoprotease which is a serine or a metallo-endoprotease or a combination of a serine and a metallo-endoprotease.

16. A method according to anyone of claims 13 to 15 wherein the protein hydrolysate is recovered without an ultrafiltration or a microfiltration step.

17. A method according to anyone of claims 13 to 16, wherein the protein substrate in incubated with an enzyme composition comprising a proline-specific endoprotease, subtilisin, and carboxypeptidase.

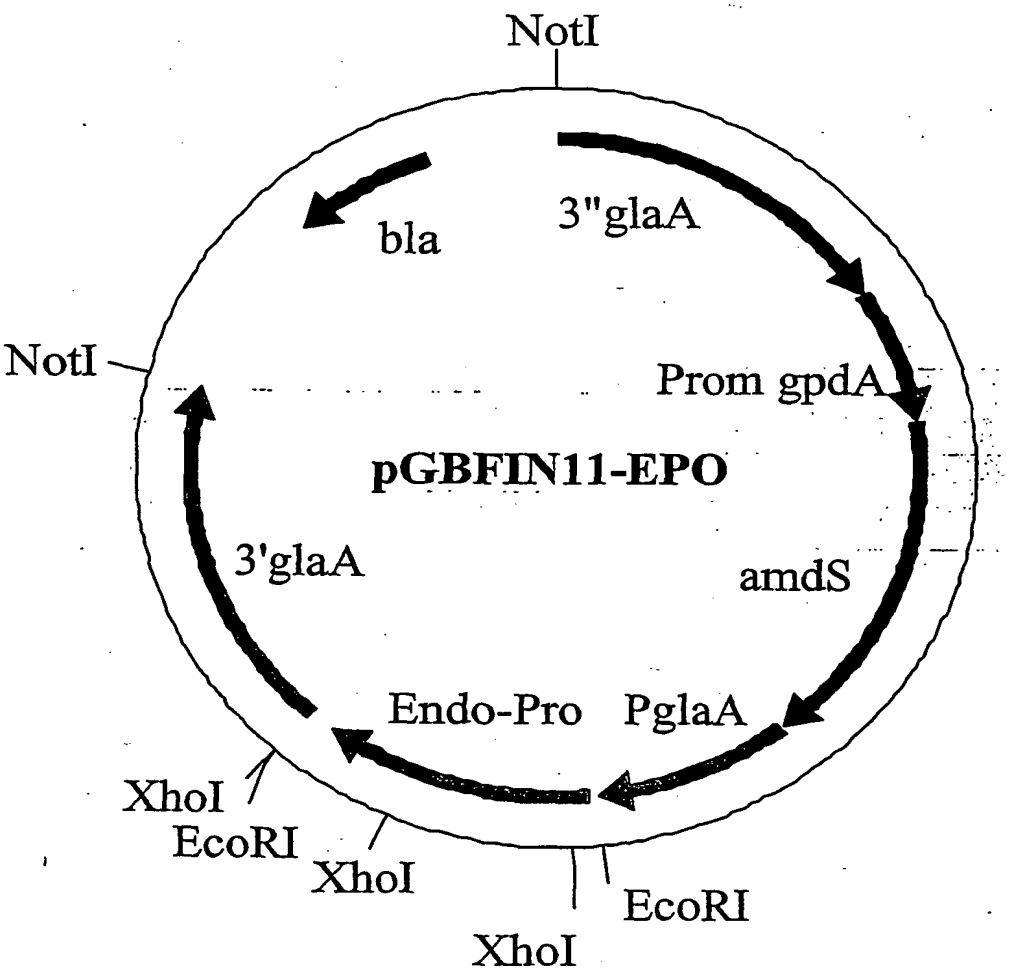

Figure 1

SEQUENCE LISTING

<110> DSM NV

<120> PROTEIN HYDROLYSATES ENRICHED IN PEPTIDES HAVING A CARBOXY TERMINAL PROLINE RESIDUE

<130> 20001WO

<160> 17

<170> PatentIn version 3.1

<210> 1
<211> 1581
<212> DNA
<213> Aspergillus niger G306

<400> 1
atgcgtgcct tctccgctgt cgctgctgcg gccctggcgc tctcttgggc gtctctggct 60

caggctgctc gcccccgtct tgtgcccaag cctgtctctc ggccagcctc gagtaaatcg 120

gctgcgacca cgggcgaggc ttactttgag cagctgctgg accatcataa tccggagaag 180

ggcacctttt cccagaggta ctggtggagt actgaatact ggggtggtcc tgggtcaccg 240

gtcgtcctct ttactcctgg agaggtctct gccgatggct atgaggggta tctcaccaat 300

gggactctca ctggtgtcta tgcgcaggag atccagggtg ccgtcattct cattgagcac 360

cgctactggg gtgattcttc tccttatgag gtgctcaatg ccgaaactct tcagtacctc 420

acactggacc aagccattct ggacatgacc tacttcgccg agacggtgaa gctgcaattc 480

gataacagca cccgcagcaa tgcgcagaat gctccctggg tcatggtcgg tggatcatac 540

agtggtgcct tgacggcttg gaccgaatct gtcgcgcctg gaacgttctg ggcttaccat 600

gccactagtg ctcctgtgga ggctatctac gactattggc aatactttta ccccatccag 660

caaggtatgg cacagaactg cagcaaggac gtgtctctgg tagccgagta tgtcgacaag 720

attggaaaga acggaactgc caaggagcag caggcactca aggaattgtt tggtctggga 780

gctgttgagc attttgatga ctttgccgct gtcctcccca acggaccgta cctctggcaa 840

gacaacgact ttgccacggg atactcttcc ttcttccagt tctgtgacgc cgtcgagggt 900

gtcgaagccg gcgcggcagt aacccccggc cccgagggtg tcggcctcga aaaggccctg 960

gccaactacg caaactggtt caattcaacc attctccctg attactgcgc aagctacggc 1020

tactggaccg acgaatggag cgtcgcctgc ttcgacagct acaacgcctc gagccccatc 1080

tacaccgata cctccgtagg caatgccgtc gaccgccaat gggaatggtt cctctgcaac 1140

130 135 140

Ala Ile Leu Asp Met Thr Tyr Phe Ala Glu Thr Val Lys Leu Gln Phe
145 150 155 160

Asp Asn Ser Thr Arg Ser Asn Ala Gln Asn Ala Pro Trp Val Met Val
165 170 175

Gly Gly Ser Tyr Ser Gly Ala Leu Thr Ala Trp Thr Glu Ser Val Ala
180 185 190

Pro Gly Thr Phe Trp Ala Tyr His Ala Thr Ser Ala Pro Val Glu Ala
195 200 205

Ile Tyr Asp Tyr Trp Gln Tyr Phe Tyr Pro Ile Gln Gln Gly Met Ala
210 215 220

Gln Asn Cys Ser Lys Asp Val Ser Leu Val Ala Glu Tyr Val Asp Lys
225 230 235 240

Ile Gly Lys Asn Gly Thr Ala Lys Glu Gln Gln Ala Leu Lys Glu Leu
245 250 255

Phe Gly Leu Gly Ala Val Glu His Phe Asp Asp Phe Ala Ala Val Leu
260 265 270

Pro Asn Gly Pro Tyr Leu Trp Gln Asp Asn Asp Phe Ala Thr Gly Tyr
275 280 285

Ser Ser Phe Phe Gln Phe Cys Asp Ala Val Glu Gly Val Glu Ala Gly
290 295 300

Ala Ala Val Thr Pro Gly Pro Glu Gly Val Gly Leu Glu Lys Ala Leu
305 310 315 320

Ala Asn Tyr Ala Asn Trp Phe Asn Ser Thr Ile Leu Pro Asp Tyr Cys
325 330 335

Ala Ser Tyr Gly Tyr Trp Thr Asp Glu Trp Ser Val Ala Cys Phe Asp
340 345 350

Ser Tyr Asn Ala Ser Ser Pro Ile Tyr Thr Asp Thr Ser Val Gly Asn
355 360 365

<220>
<221> modified_base
<222> (1)..(26)
<223> /mod_base="inosine" at positions 3,6,9,12 and 18

<400> 4
gcaacaacag gagargcata yttyga 26

<210> 5
<211> 26
<212> DNA
<213> synthetic construct
<220>
<221> modified_base
<222> (1)..(26)
<223> /mod_base="inosine" at positions 9,15,18,21 and 24

<400> 5
tcraartaag cytcaccagt agtagc 26

<210> 6
<211> 14
<212> PRT
<213> Aspergillus niger G306

<400> 6

Ala Thr Val Asn Ser Trp Thr Gly Gly Trp Asp Phe Thr Arg
1 5 10

<210> 7
<211> 23
<212> DNA
<213> synthetic construct

<220>
<221> modified_base
<222> (1)..(23)
<223> /mod_base="inosine" at positions 6,9 and 12

<400> 7
tggacaggag gatgggaytt yac 23

<210> 8
<211> 23
<212> DNA
<213> synthetic construct

<220>

Glu Arg Glu Ala Gly Ala Ala Val Thr Pro
1 5 10

<210> 13
<211> 23
<212> DNA
<213> synthetic construct

<220>
<221> modified_base
<222> (1)..(23)
<223> /mod_base="inosine" at positions 6,9,12,15,18 and 21

<400> 13
gargcaggag cagcagtaac acc 23

<210> 14
<211> 23
<212> DNA
<213> synthetic construct

<220>
<221> modified_base
<222> (1)..(23)
<223> /mod_base="inosine" at positions 3,6,9,12,15 and 18

<400> 14
ggagtaacag cagcaccagc ytc 23

<210> 15
<211> 3290
<212> DNA
<213> Aspergillus niger CBS513.88

<400> 15
gagaggcaga aggagtcatt tatcacttgt attccaatgt attttccatt tatagatact 60

gcattcaaat gcaccgttta gcatagcatc ccacattcta tttcattcca atctcatgcc 120

attgccatcc ccggtattaa tttacttctc cgccttatct tgcaatcttg caatctcttt 180

ctcctcgtta tcacgcgttc ctgcaggcgc acctccgatg gcactgcagc cggagtcccc 240

gcggcgccgg cactactaaa gactaaagtg tctagtctag cctccaatgt gctcacctcc 300

atcagcatct catccattta tcttctgacg atgtcatctg caggctccac cccctccggc 360

cgccccgacg ctctccgacg gtgcacaaca atcaattctg cagtcacgct caagattcgt 420

ccctgccgga ctcctcatgc cgtgcctggt ttaatctatg caatggagta aggtagtatc 480

gcctagcagg agcggagttc ctgctgcgct cacgccatgg tgccggcgca gacataaatc 540

taacacacga acaattccac taacaaagat ccagcggtgc ccccgaggga acctccacta 2280

ttgtgccccg gctcgtcagc gcctcctact ggcaacgcca atgcccgctc tacttccccg 2340

aagttaacgg ctacacgtac ggcagcgcga agggtaaaaa ctccgctacg gtgaacagct 2400

ggacgggtgg atgggatatg acccgcaaca cgacgcggtt gatctggacg aacgggtagg 2460

tctcccccta atttccgttg aatgtgatgt gaagataaac tcaatgctaa taaattgaga 2520

aggcaatatg acccctggcg cgactccggt gtgtcgagca ctttccggcc cggtggtccg 2580

ctggttagca cggcgaacga acccgtgcag attattccgg gcgggttcca ttgctcggac 2640

ttgtatatgg aggattacta tgcgaatgag ggtgtgagga aggtggttga taatgaggtg 2700

aagcagatta aggagtgggt ggaggagtat tatgcttgat gaagatactg gtggacatat 2760

ggagtgtaca taagatgaat ggtcataaaa tgatgatggt agatacggct atggctgttg 2820

attagatggt cctttcgcat ttcctaatta ctgagcacgt gctccatggt atgggaagtg 2880

gagacgttgc tatatatatt gactgtcggg ctattgttca cggcgtagaa gctagacgct 2940

ttgtctatgt ggccttcact aaagaccgtg actctgccca gtcttccccc cttcgaggac 3000

ctggtattag ccaaacccac ccacaaacct aacaaagatc atcgtgacat tgaagtcact 3060

ctaggtactg ctggcgctga ttacagtggc tcaattcgaa catttcaaca gcacataagg 3120

gaagggtcgc ttcacttgct accttgatac gaaagcagcc acgcccaaca cttatagggg 3180

tgacaaccat cggcatgctg ggttatctac tatatctcct gattctgtgg atcctggaga 3240

tcgatctggt acactaatct actacaatgc atgtgaagta gggataggca 3290

<210> 16
<211> 1581
<212> DNA
<213> Aspergillus niger CBS513.88

<400> 16
atgcgttcct tctccgttgt cgctgccgcg tcactggcgc tctcttgggc gtctctggcc 60

caggctgctc gcccccgtct tgtgcccaag cctatctctc ggccagcttc gagtaagtcg 120

gctgcgacta cgggtgaggc ttattttgag cagctgctgg accatcacaa cccggagaag 180

ggaacgtttt cccagcggta ctggtggagt actgaatact ggggtggacc tgggtcaccg 240

gtggtcctct ttaaccctgg agaggtctct gccgatggct atgaggggta tctcaccaac 300

gatactctca ctggtgtcta tgcgcaggag atccagggtg ccgtcattct cattgaacac 360

cgctactggg gcgactcttc gccttatgag gtgctcaatg ccgaaacact tcagtatctc 420

35 40 45

Phe Glu Gln Leu Leu Asp His His Asn Pro Glu Lys Gly Thr Phe Ser
50 55 60

Gln Arg Tyr Trp Trp Ser Thr Glu Tyr Trp Gly Gly Pro Gly Ser Pro
65 70 75 80

Val Val Leu Phe Asn Pro Gly Glu Val Ser Ala Asp Gly Tyr Glu Gly
85 90 95

Tyr Leu Thr Asn Asp Thr Leu Thr Gly Val Tyr Ala Gln Glu Ile Gln
100 105 110

Gly Ala Val Ile Leu Ile Glu His Arg Tyr Trp Gly Asp Ser Ser Pro
115 120 125

Tyr Glu Val Leu Asn Ala Glu Thr Leu Gln Tyr Leu Thr Leu Asp Gln
130 135 140

Ser Ile Leu Asp Met Thr Tyr Phe Ala Glu Thr Val Lys Leu Gln Phe
145 150 155 160

Asp Asn Ser Ser Arg Ser Asn Ala Gln Asn Ala Pro Trp Val Met Val
165 170 175

Gly Gly Ser Tyr Ser Gly Ala Leu Thr Ala Trp Thr Glu Ser Ile Ala
180 185 190

Pro Gly Thr Phe Trp Ala Tyr His Ala Thr Ser Ala Pro Val Glu Ala
195 200 205

Ile Tyr Asp Phe Trp Gln Tyr Phe Tyr Pro Ile Gln Gln Gly Met Ala
210 215 220

Gln Asn Cys Ser Lys Asp Val Ser Leu Val Ala Glu Tyr Val Asp Lys
225 230 235 240

Ile Gly Lys Asn Gly Thr Ala Lys Glu Gln Gln Glu Leu Lys Glu Leu
245 250 255

Phe Gly Leu Gly Ala Val Glu His Tyr Asp Asp Phe Ala Ala Val Leu
260 265 270

Met Glu Asp Tyr Tyr Ala Asn Glu Gly Val Arg Lys Val Val Asp Asn
500 505 510

Glu Val Lys Gln Ile Lys Glu Trp Val Glu Glu Tyr Tyr Ala
515 520 525

-13-

-1-

-1-

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 June 2002 (13.06.2002)

PCT

(10) International Publication Number
WO 02/045523 A3

(51) International Patent Classification[7]: A23J 3/34, 3/16, C12N 9/62, 9/48

(21) International Application Number: PCT/EP01/14479

(22) International Filing Date: 6 December 2001 (06.12.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
00204404.8 7 December 2000 (07.12.2000) EP
01204464.0 15 November 2001 (15.11.2001) EP

(71) Applicant *(for all designated States except US)*: **DSM N.V.** [NL/NL]; Het Overloon 1, NL-6411 Te Heerlen (NL).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **EDENS, Luppo** [NL/NL]; Hoflaan 118, NL-3062 JL Rotterdam (NL). **HOEVEN, VAN DER, Robertus, Antonius, Mijndert** [NL/NL]; Kerkstraat 6a, NL-2223 AJ Katwijk (NL). **DELEST, Veronique** [FR/FR]; 53 Rue Du Chateau, F-92100 Boulogne (FR).

(74) Agents: **MATULEWICZ, Emil, Rudolf, Antoni** et al.; DSM N.V., DSM Patents & Trademarks, P.O. Box 1, NL-2600 MA Delft (NL).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
30 January 2003

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/045523 A3

(54) Title: PROTEIN HYDROLYSATES ENRICHED IN PEPTIDES HAVING A CARBOXY TERMINAL PROLINE RESIDUE

(57) Abstract: A method of enzymatically producing a protein hydrolysate from a protein substrate is described, wherein a proline-specific endoprotease or a composition containing a proline-specific endoprotease and optionally a subtilisin or a metallo endoprotease, and other enzymes such as carboxypeptidases, is used to produce a protein hydrolysate enriched in peptide fragments having a carboxy terminal proline residue. Such protein hydrolysates may be used as such or to reduce bitterness in foods nutritionally supplemented by protein hydrolysates, as well as to produce hydrolysate-containing foodstuffs having low antigenicity.

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PATENT ABSTRACTS OF JAPAN vol. 017, no. 278 (C-1065), 28 May 1993 (1993-05-28) & JP 05 015314 A (FUJI OIL CO LTD), 26 January 1993 (1993-01-26) cited in the application abstract	9-11,24
A	PATENT ABSTRACTS OF JAPAN vol. 1997, no. 05, 30 May 1997 (1997-05-30) & JP 09 000164 A (ASAHI CHEM IND CO LTD), 7 January 1997 (1997-01-07) abstract	1-17
A	DATABASE WPI Week 199249, 21 October 1992 (1992-10-21) Derwent Publications Ltd., London, GB; AN 1992-401807 XP002166608 & JP 04 297923 A (AJINOMOTO KK) abstract	1-17
A	PATENT ABSTRACTS OF JAPAN vol. 014, no. 197 (C-0712), 23 April 1990 (1990-04-23) & JP 02 039896 A (EZAKI GLYCO KK), 8 February 1990 (1990-02-08) abstract	1-20